WPS4286

POLICY RESEARCH WORKING PAPER 4286

Institutional Effects as Determinants of Learning Outcomes

Exploring State Variations in Mexico

Jesús Álvarez
Vicente García Moreno
Harry Anthony Patrinos

The World Bank
Human Development Network
Education Team
2007

Policy Research Working Paper 4286

Abstract

This paper uses the OECD's Program for International Student Assessment student-level achievement database for Mexico to estimate state education production functions, controlling for student characteristics, family background, home inputs, resources, and institutions. The authors take advantage of the state-level variation and representative sample to analyze the impact of institutional factors such as state accountability systems and the role of teachers' unions in student achievement. They argue that accountability, through increased use of state assessments, will improve learning outcomes. The authors also cast light on the role of teachers' unions, namely their strength through appointments to the school and relations with state governments. The analysis shows the importance of good relations between states and unions. Furthermore, it demonstrates that accountability systems are cost-effective measures for improving outcomes.

This paper—a product of the Education Team, Human Development Network—is part of a larger effort in the network to analyze the determinants of learning. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Shaista Baksh, room G8-056, telephone 202-473-1085, fax 202-522-3233, email address Sbaksh@worldbank.org. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. Harry Patrinos may be contacted at hpatrinos@worldbank.org. July 2007. (24 pages)

INSTITUTIONAL EFFECTS AS DETERMINANTS OF LEARNING OUTCOMES:

EXPLORING STATE VARIATIONS IN MEXICO

Jesús Álvarez
Vicente García Moreno
Harry Anthony Patrinos*

JEL Codes: I2, J24, H52, L33

Keywords: Student assessment, education outcomes, Mexico, accountability, unions

* Background paper prepared for the Mexico/World Bank Study on the Quality of Education in Mexico. The views expressed here are those of the authors and should not be attributed to the World Bank Group. We thank April Harding and participants at seminars in Mexico City and Washington DC for useful comments. Address all correspondence to Harry Patrinos (hpatrinos@worldbank.org).

Introduction

Previous research confirms the importance of socioeconomic status on learning and the limited role of physical investments (see, for example, World Bank 2005). It is also expected that school climate, expectations, participation, autonomy, accountability and the use of assessments will have a significant impact on learning outcomes. It is also expected that an education system that is based on constant assessment and participation in international benchmarking exercises will improve its effectiveness. In most of the countries that performed well on the Organisation for Economic Cooperation and Development's (OECD) Programme for International Student Assessment (PISA), local authorities and schools have substantial responsibility for educational content and/or the use of resources, and many set out to teach heterogeneous groups of learners (see, for example, Fuchs and Woessmann 2006).

Mexico has been participating in PISA since its inception in 2000. This marked a significant change in the use of assessments and transparency in Mexico, where results were previously not made publicly available. Mexico's scores on PISA are below average, but no worse than for other countries in Latin America, except for Uruguay in 2003, but in all cases Mexico shows a lower level of inequality in test scores than all other Latin American participating countries. For Mexico, there has not been much improvement since PISA 2000. In PISA 2003, Mexico's performance in all three subjects (mathematics, science and reading) declined, though this may be associated with the fact that enrollments increased during the same period of time by about 5 percentage points. The Government of Mexico (2005), the OECD (2005), and the World Bank (2005) call for broader use of results to influence policy decisions, school management and users' choice.

In this paper we take advantage of the fact that Mexican data are representative at the state level to include more variables at the state level. This is done in an effort to measure the importance of state accountability systems, decentralization and union power on student learning outcomes. The analysis reaffirms the importance of school climate, but also supports the contention that further decentralization, school autonomy and assessment are important for improving learning outcomes in Mexico. It also points to the fact that the states are able to align their policies to ensure that what works at the local level materializes.

Review

Researchers have begun to use international assessments to analyze the determinants of learning. Hanushek and Luque (2003) indicate that attention to the quality of human capital in different countries naturally leads to concerns about how school policies relate to student performance. Using the Third International Mathematics and Science Study (TIMSS), the results of their analyses of the educational production functions within a range of developed and developing countries show general problems with efficiency of resource usage similar to those found previously in the United States. These effects did not appear to be dictated by variations related to income level of the country or level of resources in the schools, and the conventional view that school resources are relatively more important in poor countries also failed to be supported.

At the country level some research using international assessments has appeared. Fertig (2003) used OLS and quantile regressions to analyze the determinants of German students' achievement using PISA 2000. Among the negative suggested factors were: schools without regular tests; too much regulation of schools; poor school conditions; not enough access to modern information technology for the students; non-native students; and high student-teacher ratio and shortage of teachers. Fertig and Schmidt (2002) provided, based on the individual-level data of the PISA 2000 study, a detailed econometric analysis of the way that reading test scores are associated with individual and family background information and with characteristics of the school and class of the 15 year old respondents to the survey. Based on quantile regressions, they interpreted the national performance scores, conditional on these observable characteristics, as the reflection of different education systems. Their findings suggest that United States students, particularly those in the lower quantiles, are served relatively unsatisfactorily by their system of education. Wolter and Vellacott (2002) analyzed the sibling size and birth-order effect on educational achievement in Switzerland on the basis of PISA data. They show that, besides the usual factors like education, wealth or the occupational status of parents, family configurations can play an important role in explaining differences between students.

Countries around the world are moving toward increased accountability of schools for student performance. The United Kingdom has an elaborate system of league tables giving

parents information about the performance of schools in terms of test scores and other indicators. The United States has legislated that all states develop an accountability system. Evidence on the impacts of these systems is growing. United States evidence indicates that strong accountability systems lead to better student performance (Carnoy and Loeb 2002; Hanushek and Raymond 2005; Jacob 2005). Less evidence is available about accountability systems in developing countries. This could be due to weak accountability in these countries, along with a general lack of systematic measurement and reporting of student achievement.

In an important paper, Woessmann (2003), using TIMSS, suggests that international differences in educational institutions explain the large international differences in student performance in cognitive achievement tests. An econometric student-level estimation based on data for more than 260,000 students from 39 countries reveals that positive effects on student performance stem from centralized examinations and control mechanisms, school autonomy in personnel and process decisions, competition from private educational institutions, scrutiny of achievement, and teacher influence on teaching methods. A large influence of teachers' unions on curriculum scope has negative effects on student performance. The findings imply that international differences in student performance are not caused by differences in schooling resources but are mainly due to differences in educational institutions. Taking all countries into consideration, he finds that the following factors positively impact science and mathematics learning: central examinations; centralized control of curriculum and budget matters; school autonomy in process and personnel; teacher incentives; limited influence of unions; scrutiny of student performance; parental interest; intermediate level of administration; and competition from the private sector. Fuchs and Woessmann (2006) obtain similar results using PISA 2000. In fact, they find that 25 percent of the variation in scores is attributable to institutional variation. Student performance is higher with external exams and budget formulation, but also with school autonomy in textbook choice, hiring teachers and within-school budget allocations. School autonomy is more beneficial in systems with external exit exams.

It is argued that teachers' unions may have a negative impact on learning outcomes (Hoxby 1996; Woessmann 2003). Moreover, in Mexico, the main teachers' union (*Sindicato Nacional de Trabajadores de la Educación*, or National Union of Education Workers, or SNTE)

is large, powerful and well organized. It was established in 1943, interestingly enough by the then Secretary of Public Education, Jaime Torres Bodet (Ornelas 1988), who went on to become Secretary General of UNESCO from 1948 to 1952. SNTE was created as a very centralized and monopolistic organization, formed from the merger of Union of Education Workers (SUNTE), the Mexican Union of Teachers and Education Workers (SMMTE), the Autonomous National Union of Education Workers (SNATE), and the Union of Workers of Mexican Education (STERM), as well as other smaller groups (Murillo 1999). While there are other unions, SNTE is the largest, with 1.4 million members. Until 1992 it was affiliated with the longtime incumbent Institutional Revolutionary Party (PRI), serving a political role for a long time, especially during elections. Teachers' demonstrations are frequent occurrences. All public school teachers in Mexico belong to a teachers' union, but not by choice. While it could play a critical role in improving quality, it has so far given priority to raising members' salaries and expanding teaching staff. Recently the teachers' union has become more active in political issues, this time free of any political party affiliation. Some argue that the teachers' unions are a barrier to reform and improvement of the education system in Mexico (Ornelas 2004).

Overall union density has gone down in Mexico since 1984, from 30 to 21 percent, and this includes teachers (Fairris and Levine 2004). There was a decline in the proportion of education sector workers (not just teachers, but also administrators, secretariat staff, etc.) that are unionized, from 73 percent in 1984 to 64 percent in 2000; still, teachers remain the most unionized segment of the labor force. In fact, all public school teachers belong to a union; it is mandatory. This is a higher proportion than Korea (5 percent), Singapore (22 percent), Great Britain (60 percent), Spain (63 percent), the United States (68 percent), the Netherlands (80 percent), Canada (81 percent) and Denmark (95 percent) (Kasten and Fossedal n.d.).

Another measure for union power is the level of conflict that exists between the state and the teachers' union. Unfortunately, in Mexico there is no official central registry of number of days that schools are closed due to strike activity. In fact, days away from school during strikes are not counted as teacher absenteeism. Conflict could be said to be the result of a lack of political alignment due to lack of trust and coordination problems that make negotiations difficult. Conflict between the state and the teachers' union was used by Murillo and others

(2002) in a study for Argentina. Conflict is found to have a negative effect on learning outcomes in Argentina (Murillo and others 2002). Adversarial political alignments are associated with a decrease in effective numbers of class days, with an indirect negative effect on student performance in Argentina. A recent survey for the Latin America region finds that strike activity by Mexican teachers is one of the highest in the region. Between 1998 and 2003, there were 49 strikes in Mexico; much more than in Chile (4) or Costa Rica (5), but much less than in Argentina (93) or Brazil (90). The strikes in Mexico led to 434 lost days of schooling throughout the country (Gentili and Suarez 2004).

Methodology

We analyze the determinants of school achievement in Mexico using ordinary and generalized least squares. Factors affecting achievement are analyzed and compared. In this regard ordinary least squares (OLS) methods are used to analyze the determinants of learning. The following linear regression model is estimated:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon \qquad (1)$$

where Y is the test score and X_1 is a vector of student variables that include household characteristics such as socioeconomic indicators, and X_2 is a vector of school indicators such as school resources, school and institutional features. It is expected that the scores among students in the same schools will be correlated. The reason is that students enrolled in the same school are usually more similar to one another in behavior and characteristics than students enrolled in different schools. In other words, one would expect that student performance for given school factors would increase in order for those school variables to increase or improve, but one might also expect the variation on average school performance to increase as school factors increase or improve. However, because of the non-spherical error term ($\varepsilon \approx N(0, \sigma^2\Omega)$), the OLS estimation is not thought to be highly dependable. The OLS estimate does not account for dependency due to clustering effects. Other OLS estimates take into account the sampling procedure, but the correlation between other school characteristics implicit in the survey (location, type, level and program) would not be corrected. In order to accommodate for schools fixed effects we use the generalized least squares (GLS) estimation methodology. To

accommodate the school factors and cover for the between schools and within schools dimensions we estimate a combined model:

$$Y = \beta_1 X + \mu S + \varepsilon \qquad (2)$$

where X is the predictors' matrix that also includes the school and institutional variables – which are fixed for each student at the same school; S is the predictors matrix that includes student variables only; μ is a random element associated with school disturbances (as a second level random variable), which we assume to have covariance matrix T. We use the GLS estimate for β as $\beta^* = (X' V^{-1} X)^{-1} X' V^{-1} Y$, where V is the variance matrix and is equal to $ZT^d Z' + \sigma^2 I$, and T^d is the diagonal matrix for the variance of μ. Since T and σ^2 are most likely to be unknown we estimate their values to fit the parameters by GLS. For the estimation, iterative generalized least squares will be used.

Thus, we use the same basic model as in World Bank (2005), but add new institutional variables that were recently collected for each state. On the modeling of institutional variables in education production functions, see Bishop and Woessmann (2004). This allows us to see how state authorities' actions affect learning outcomes. More specifically, we use PISA 2003 to estimate the determinants of learning outcomes, and take advantage of the fact that the Mexican data are representative at the state level and by type of school. Test scores, household and socioeconomic status variables are obtained at the student level, while resources and institutional features surrounding students' learning are measured at the classroom, school and state level.

Data

The student population in PISA is 15 year-olds, who are thus assessed as they approach the end of their compulsory schooling. For more information about the design, development and implementation of PISA, see http://www.pisa.oecd.org. Mexico was the only country that expanded the sample to include state representatives with a random sample of 29,983 students chosen from 1124 schools that participated in the assessment from all states (except Michoacan) and the Federal District. The survey was carried out in two stages; the explicit stratification was based on states and size of the schools, the implicit stratification was based on school type, urban/rural, school level and school program. Because the survey comprises three different

questionnaires (cognitive skills, student and school questionnaires), there are variables with missing information for some students.

Table 1: Descriptive statistics of Variables used in the Analysis

	Mean	s.d.
Scores		
Math	385.3	80.6
Science	405.0	76.4
Reading	399.8	86.5
Student characteristics		
Town less than 15,000 (%)	0.350	
City less than 1,000,000 (%)	0.499	
City more than 1,000,000 (%)	0.151	
Age (years)	15.80	0.3
Female	0.500	
Family Background		
Mother with lower secondary complete (%)	0.592	
Mother working (%)	0.350	
Home incentives and inputs		
Homework (hours)	6.9	5.9
Home educational resources (index)	-0.5	1.2
Internet (index)	3.1	1.8
Use of computer at home (index)	3.4	1.5
School resources		
Motivation in Math (index)	0.6	0.6
Memorization (index)	0.5	1.0
Teacher Morale (index)	0.01	1.1
Sense of belonging to school (index)	0.2	1.0
Private school (%)	0.6	
Girls in school (%)	0.5	

Source: PISA 2003

* Using item response theory, PISA mapped performance in each subject on a scale with an international mean of 500 test-score points across all OECD countries and an international standard deviation of 100 test-score points across the OECD countries

We excluded all student observations from the analysis that have a missing value of at least one variable. The learning domains of reading, mathematical and scientific literacy,

together with some other areas such as students' familiarity with computers, learning strategies, and students' attitudes towards their schools, have been chosen to be the foci of PISA. PISA's assessment materials focus on young people's ability to apply their knowledge and skills to real-life problems and situations, rather than on how much curriculum-based knowledge they possess. The emphasis is on whether students, faced with problem situations that might occur in real life, are able to analyze, reason and communicate their ideas, arguments or conclusions effectively. The term literacy is attached to each domain to reflect the focus on these broader skills. In the way that the term is used, it means much more than the traditional meaning of being able to read and write. The variables used in the analysis are listed in Table 1.

A number of institutional variables were included in the analysis, taking advantage of the fact that Mexican data from PISA 2003 are fully representative at the state level. These new variables, therefore, are measured at the state level (Table 2; see also Annex Table 1).

Table 2: Institutional Variables Means and Definitions

Variable	Mean (s.d.)	Variable range	Definition
Administrative decentralization (within state)	0.50 (0.5)	**0-1**	**State oversight of administrative issues has been moved from the state capital to the municipal level**
Pedagogical decentralization (within state)	0.20 (0.4)	**0-1**	**State has allowed pedagogical decision-making to vary by locality**
State evaluation system		**1-5**	**Level of evaluation state implements:**
1st stage	0.26 (0.4)	1	Only national evaluations
2nd stage	0.34 (0.5)	2	States have own tests
3rd stage	0.13 (0.3)	3	States disseminate results
4th stage	0.20 (0.4)	4	States receive feedback from schools
5th stage	0.07 (0.3)	5	Sates design policy, strategy, interventions
Union power		**1-3**	**Level of teachers' union influence on teacher appointment:**
	0.07 (0.3)	1	Low
	0.45 (0.5)	2	Medium
	0.49 (0.5)	3	High
Conflict		**1-3**	**Level of conflict between state government and teachers' union:**
	0.62 (0.5)	1	No significant conflict
	0.08 (0.3)	2	Exist conflict
	0.30 (0.5)	3	High conflict

We introduce variables describing within-state decentralization. Both are 0-1 dummy variables indicating whether or not the decentralization took place. There are two such variables: (1) administrative decentralization—moving state oversight from the state capital to the municipal level and (2) pedagogical decentralization—allowing decision making to vary by locality (for example, capacity of schools to define training needs, capacity of zone supervisors to jointly develop with schools improvement plans, capacity of regional offices to develop programs of academic improvement based on test scores). Such actions, it could be argued, may have been allowed in order to put people at the center of service provision since it is believed that can go a long way towards improving service delivery. Focusing on people enables them to monitor and discipline service providers and amplifies their voice in policymaking, and strengthens the incentives for providers to serve them (World Bank 2004). The states that have

decentralized the pedagogical functions have brought key decision-making closer to the school and beneficiaries. Twenty percent of Mexican states have done this. By contrast, half of all states have decentralized administration within the state.

Accountability systems – student testing, school rankings, school report cards – are believed to have a strong impact on improving service delivery, thus making them good candidates for improving learning outcomes (see, for example, World Bank 2004). We developed five categories of state accountability systems: (1) states that rely only on important yet sample survey national student assessments carried out by a national agency on behalf of the national government (that is, they do not implement, report on or use state-level examinations)—26 percent have only this; (2) states that do not only rely on national assessments, but implement their own examinations of students in their schools (34 percent of states); (3) states that use their state-wide assessments systems to inform the public by, for example, disseminating results to the school (13 percent); (4) states that received feedback on the results from the schools (20 percent); and (5) states that use the results and the public feedback to design policies, strategies and specific interventions to improve outcomes (7 percent). The fifth level is what we consider the complete or full accountability state system. It is believed that accountability systems could be particularly useful investments if they contribute to improved learning outcomes, especially given their extremely low cost (see Hoxby 2002).

In this study, we have information on the power of unions – given that all public school teachers are unionized one cannot identify states with and without unions, nor can we in any way replicate the seminal study by Hoxby (1996) who used differences in the timing of collective bargaining agreements across states in the United States, nor look at the impact of union density or fragmentation (as Murillo and others 2002 did for Argentina). Our information on teacher union power ranges from low in terms of influencing the allocation of teacher positions, to medium, and high. High would refer to states where the unions allocate all teachers—this characterizes 50 percent of Mexican states; medium refers to states where 50 percent of allocations are made by the union and 50 percent through competitive examinations managed by state authorities (about 45 percent); and low refers to states where unions allocate less than 50 percent of teachers (only 7 percent).

Another measure for union power is the level of conflict that exists between the state authorities and the teachers' union in that state. Our conflict variable is constructed through state officials contacted in each case by the same person, one of the co-authors of this paper, who interviewed state officials and elicited responses to a question about the frequency and seriousness of disagreements between state authorities and the section of the union represented in the state since 2000. The conflict variable is categorized as follows: (1) disagreements exist, but they are not serious (62 percent of states); (2) the disagreements are frequent but not profound; they are manifested in declarations in the mass media (8 percent); and (3) almost every year there are profound disagreements; they are manifested in marches, taking over facilities and, in many occasions, suspension of school activities (30 percent). Murillo and others (2002) use a similar variable in Argentina. It also conforms to the situation described in Grindle (2004) and Ornelas (n.d.) in terms of union-state relations post-1992 decentralization. Conflict could be said to be the result of a lack of political alignment due to lack of trust and coordination problems that make negotiations difficult.

Results

The full regression results are presented in Annex Tables 2 and 3. In Annex Table 2 we enter each of the institutional state-level variables one at a time. First, it is shown that further decentralization within the state has a positive, but insignificant effect. Accountability systems – student testing, school rankings, school report cards – are shown to have a strong, positive and significant impact on learning outcomes. That is, states that do not rely only on important yet sample survey national student assessments have higher scores on PISA, controlling for everything else (second stage accountability system). Further, authorities that use the results of their state-wide assessment systems to inform the public, disseminate the results to the school, received feedback from users have a significant impact on learning outcomes. While student evaluations at the state level and evaluations systems that disseminate the results back to the school have positive and significant impacts, the greatest impact comes from more complete systems that non only use the results to inform policy and disseminate results, but also use the results to design specific interventions (fifth, or complete accountability stage), have a very large impact on learning outcomes. This makes it a particularly useful investment given its large

contribution to learning outcomes as well as the fact that it is a very cheap investment (see below).

In this study, we have information on the power of unions ranging from low in terms of influencing the allocation of teacher positions, to medium, and high. Indeed, in Mexico union influence is associated with lower test scores. In our regression analysis we enter two union power variables; both are relative to low union power. A high influence is not significant. However, medium power is significant and has a relatively large negative effect.

Another measure for union power is the level of conflict that exists between the state authorities and the teachers union in that state. The conflict variable takes values of: (1) low—disagreements exist but they are not serious; (2) medium—disagreements are frequent but not profound; and (3) high--almost every year there are profound disagreements manifested in marches and suspension of school activities. Relative to high levels of conflict, only having a low level of conflict is significantly and positively associated with learning outcomes.

Full Model

However, when we include all factors together (Table 3), it turns out that only two of the new institutional variables are significant for math: (1) using the state evaluation system to feedback to schools and design interventions and (2) conflict between the union and state. The full evaluation-feedback-design (fifth stage) system has the largest impact. None of the other variables are significant. This is a strong correlation suggesting that states can take significant actions to improve their school systems by developing and using an accountability system. Thus, institutions matter, but the most significant institutional issues are relatively low cost and under the direct control of state authorities.

This is not to say that unions are unimportant, but relative union power is not a barrier to reform when states have the willingness to develop state evaluation systems and engage in further decentralization of pedagogical matters. In some states, interesting experiments are taking place to improve quality and efficiency, reflecting successful negotiations with the local sections of the teachers' union (OECD 2005). The more successful states in terms of academic

achievement, especially PISA scores, are making improvements in the selection of teachers, in collaboration with the teachers' unions in the state.

Table 3: Institutional Effects as Determinants of Student Achievement

		Math		Reading		Science		All	
Institutional factors									
Decentralization within state:	Administrative	0.4		0.0		1.7		0.7	
	Pedagogical	3.1		0.6		4.7		2.8	
Accountability:	(2nd stage)	2.0		-0.3		0.8		0.8	
	(3rd stage)	-4.4		-1.2		-7.2	*	-4.3	*
	(4th stage)	-2.1		-3.9		-5.4	*	-3.8	
	(complete)	14.7	*	12.4	*	6.1		11.1	*
Union influence on teacher positions:	Medium	-5.5		-5.4		-4.1		-5.0	
	High	1.9		2.0		1.7		1.8	
Conflict between state and union:	Medium	4.6		3.4		6.8	*	6.8	*
	Low	9.2	*	9.0	*	9.3	*	9.3	*
Controls for:									
Student characteristics		incl.		incl.		incl.		incl.	
Family background		incl.		incl.		incl.		incl.	
Home incentives and inputs		incl.		incl.		incl.		incl.	
Log Likelihood		-68,188		-68,349		-68,152		-66,727	
Observations		12,332		12,332		12,332		12,332	

Source: Estimation with GLS using PISA 2003; institutional variables; for full results, see Annex Table 3

* Denotes significance at the 99% level

For other subjects the results largely reconfirm the findings presented in the case of math. The results for reading are almost identical to those for math. In the case of science accountability systems do not seem to be important and in one case having state testing has a negative correlation. For science outcomes only better relations with the teachers' union appears to be a significant determinant of outcomes. But when we analyze all subjects together the model seems to work. Having a complete accountability system has a strong correlation with overall test scores. Less conflict between the state and teachers' union improves overall test scores. Curiously though when we consider all three subjects together union influence on teacher positions, which was never a significant variable for any one subject alone, becomes significant. There is a negative correlation between a *medium* union influence and overall test scores. A high union influence is not significant.

In addition to the previous analysis, we have used quantile regression analysis to estimate the differential contribution of the institutional variables along the distribution of student

achievement (Table 4). Similar to the results from the full model, state authorities that use the results of their state-wide assessment systems to build a strong accountability system – inform the public, disseminate the results to the schools, and get feedback from users – have a more significant impact on learning outcomes of low performing students than for high performing students. For the students in the bottom of the distribution of achievement, institutional factors have a greater impact on their learning. Also, a low level of conflict between state authorities and the teachers' union has a significant and positive effect; medium union influence on teacher positions has a negative effect. The effects of these two union-related variables imply that low achieving students are vulnerable to union power. These results also suggest the need for more transparent and accountable educational institutions in order to address the needs of disadvantaged students, as well as a better relationship between state authorities and the teachers' union.

In order to attempt to address the causality issue, given the non-experimental nature of our data, we are using a propensity score matching algorithm that identifies comparable students with similar backgrounds, but that differs in terms of exposure to state accountability systems. We are using the scores to match students of three similar states, Colima, Guanajuato and Tlaxcala. One state has a full accountability system (Colima), another one is at the mid-range of such a system (Guanajuato), and one lacks a state evaluation system (Tlaxcala). We have analyzed differences in estimated test scores based on exposure to different institutional factors at the state level. Annex Table 4 shows that the full accountability model – tests, publication, feedback and use for policy and strategy – produces significant differences and positive results. Comparing Colima with Tlaxcala, the results show that the latter, a state with a poor performance that does not have a full evaluation system, could reach the average level of performance among Mexican states if it introduces full accountability. And the comparison between Colima and Guanajuato shows that once Guanajuato implements a full accountability system, it will be one of the top performing states. Tlaxcala could improve by 0.35 standard deviations and Guanajuato by 0.22 standard deviations if they introduce full accountability.

Table 4: Institutional Effects as Determinants of Student Math Achievement across the Achievement Distribution

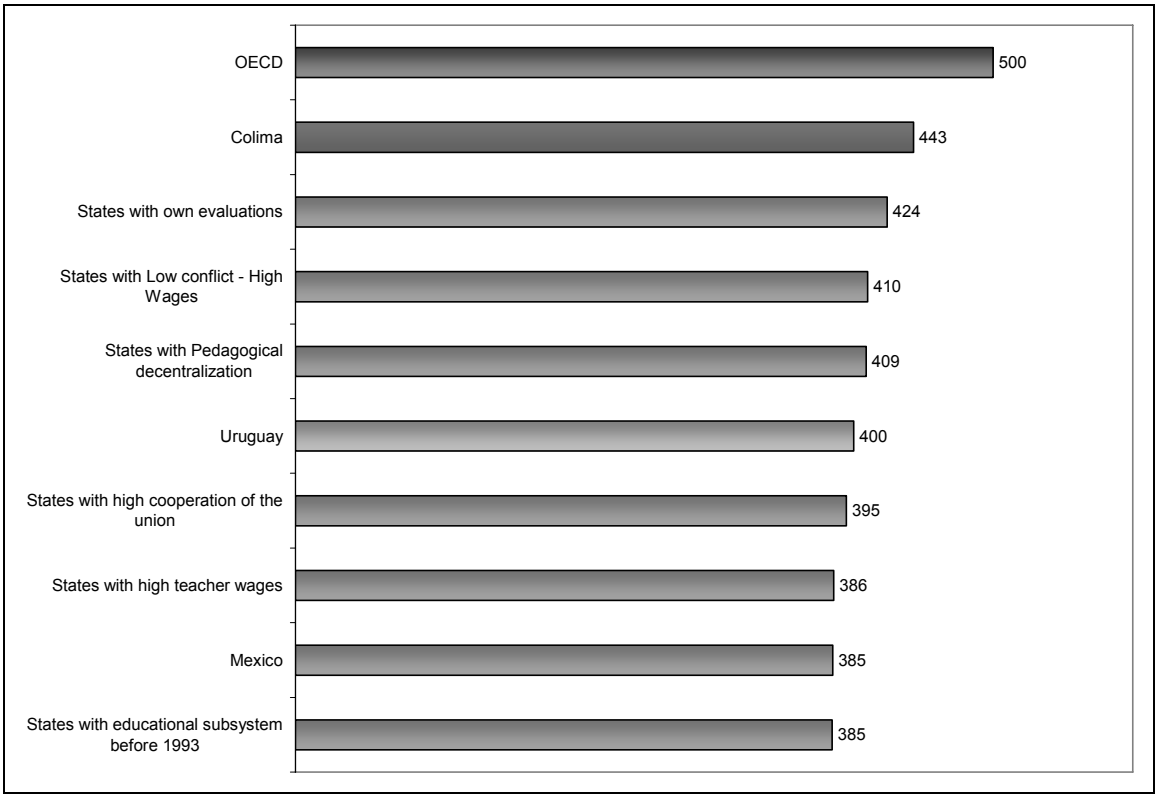
	Quantile							
Institutional factors	0.20		0.40		0.60		0.80	
Administrative decentralization	-1.7		0.3		0.6		3.8	
Pedagogical decentralization	-0.6		2.1		-1.5		3.2	
Evaluation (2^{nd} stage)	0.0		1.1		-0.8		4.6	*
Evaluation (3rd stage)	-6.4		-3.3		-2.1		-2.3	
Evaluation (4th stage)	-1.2		-3.5		-2.4		-4.3	
Evaluation (full accountability)	16.4	*	13.7	*	14.6	*	8.1	
Medium union influence teacher positions	-10.5	*	-5.7	**	-9.0	**	0.2	
High union influence teacher positions	-3.1		0.2		-2.4		4.5	
Medium conflict state and union	4.3		3.6		3.3		3.7	
Low conflict state and union	10.6	*	9.1	*	7.8	*	6.5	*
Controls for:								
Student characteristics	incl.		incl.		incl.		incl.	
Family background	incl.		incl.		incl.		incl.	
Home incentives and inputs	incl.		incl.		incl.		incl.	
Pseudo R^2	0.14		0.15		0.15		0.15	
Observations=12,332	-		-		-		-	

Source: Estimation with Quantile Regressions with Bootstrapped SE using PISA 2003; institutional variables; full results available upon request

* Denotes significance at the 99% level; **Denotes significance at 95 % level

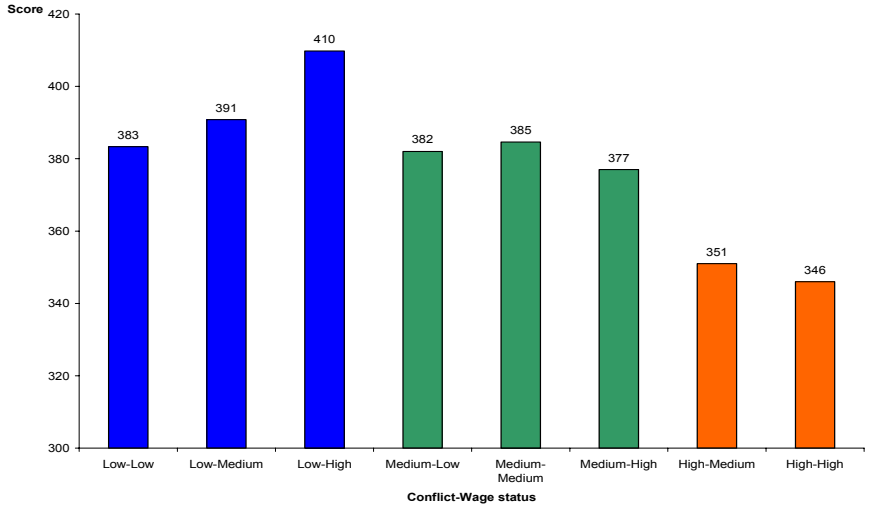
It is not enough to have low levels of conflict with unions, although it helps. More importantly, paying teachers more will not necessarily reduce conflicts, and there is no evidence that it will lead to better learning outcomes (Figure 1). States with low levels of conflict and high teacher wages do very well. Even better are states that have complete and comprehensive accountability systems. The accountability system for Colima (World Bank 2005), the best performing Mexican state, is characterized by all three factors.

Figure 1: Test Scores by Institutional Framework



It is interesting to note that a "medium" level of conflict and "medium" level of wages for teachers reproduces the exact average PISA score for math in Mexico (Figure 2). At this average level of conflict, the level of salaries is irrelevant for improving outcomes. Low salaries are not associated with good results. But low levels of conflict with high salaries appear optimal.

Figure 2: Average PISA Math Score by Teacher Wage and Union-State Conflict



Towards cost-effectiveness

The national sample-based student assessment run by the National Institute for the Evaluation of Education (INEE) is estimated to cost only $US 6 dollars per student (Table 5). This compares to other major programs such school-based management which have been evaluated to perform well (Gertler, Patrinos and Rubio-Codina 2006; Skoufias and Shapiro 2006). It also appears to be a much better investment than other, more expensive, interventions, such as high salaries for teachers or more computers. Many of the more expensive interventions are also untried or untested.

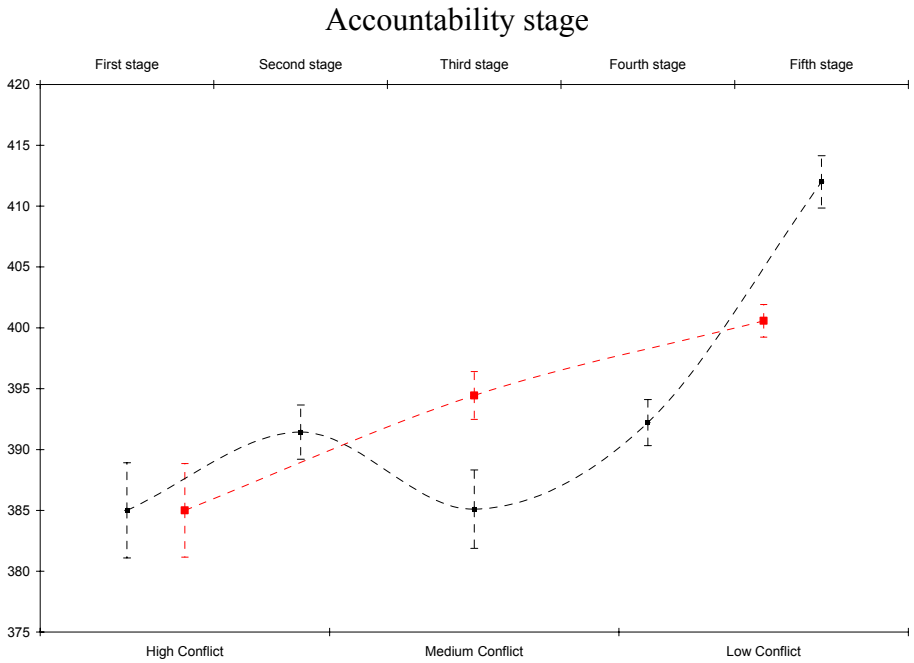
Table 5: Unit Costs of Selected Mexican Education Programs, 2005

National Student Assessment	$US 6
AGEs (*Apoyo a la Gestión Escolar*, a rural school-based management program	$US 7
State of Aguascalientes Student Assessment	$US 10
PEC (*Programa Escuelas de Calidad*, an urban school-based management program	$US 37
School building	$US 160
New teachers position and salary increase	$US 240
Computers (1 per 10 students)	$US 500
Student assessment as percentage of per pupil spending	0.70%

Note: Calculations made on the basis of a unit cost of $US 1,494 for basic education in 2005

To further assess the relative impact of accountability systems at the state level, we use the parameters produced in Table 3, and forecast PISA scores in math, controlling for everything else, and varying both (a) the level of accountability and (b) the level of conflict between the state government and the teachers' union (Figure 3). Clearly less conflict between union and government will lead to improved scores. The orders of magnitude are roughly in line with increasing levels of accountability up to the fourth stage. The increase in scores is much higher when states have full accountability systems, meaning that they implement their own assessments, use the results for policymaking, provide feedback to the schools, and use all that information to create strategies and programs.

Figure 3: Simulated Math Scores



Conclusion

The analysis of the new institutional variables suggests that, in general, more accountability (and assessment) is needed to improve learning outcomes. The analysis confirms the importance of continued use of assessments, not only at the national level for benchmarking and policy guidance, but also at the state level through universal state systems that provide constant feedback to beneficiaries and are used by the authorities to design interventions. Therefore, state-level assessments are very important. While unions will not initiative or initially support reform movements to improve the quality of education, they are important partners for gaining support for state initiatives. Much of the variation among states may be due to the priorities of governors, their perspectives on the importance of education, and the relationship they are able to build with the state teachers' unions (see also Grindle 2004). If there were only a few things that states could do to improve the quality of education, they would be to implement state accountability systems and increase school level autonomy, within a context of positive relations with the teachers' unions that would facilitate incremental reforms in the quality of teacher selection.

References

Bishop, J. and L. Woessmann. 2004. "Institutional Effects in a Simple Model of Educational Production." *Education Economics* 12(1): 17-38.

Carnoy, M. and S. Loeb. 2002. "Does external accountability affect student outcomes? A cross-state analysis." *Educational Evaluation and Policy Analysis* 24(4):305-331.

Fairris, D. and E. Levine. 2004. "Declining Union Density in Mexico, 1984-2000." *Monthly Labor Review* 127(9): 10-17.

Fertig, M. 2003. "Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study." IZA Discussion No. 739, Bonn.

Fertig, M. and C.M. Schmidt. 2002. "The Role of Background Factors for Reading Literacy: Straight National Scores in the PISA 2000 Study." IZA Discussion No. 545, Bonn.

Fuchs, T. and L. Woessmann. 2006. "What Accounts for International Differences in Student Performance? A Re-examination using PISA Data." *Empirical Economics* (forthcoming).

Gentili, P. and D. Suarez. 2004. "La Conflictividad Educativa en America Latina." Foro Latinoamericano de Politicas Educativas, Rio de Janeiro and Buenos Aires (unpublished paper).

Gertler, P., H.A. Patrinos and M. Rubio-Codina. 2006. "Empowering Parents to Improve Education: Evidence from Rural Mexico." World Bank Policy Research Working Paper No. 3935.

Government of Mexico. 2005. *Quinto Informe de Gobierno*. Presidencia de la Republica.

Grindle, M.S. 2004. "Interests, Institutions, and Reformers: The Politics of Education Decentralization in Mexico," in R. Kaufman and J. Nelson, eds, *Crucial Needs, Weak Incentives: Social Sector Reform, Democratization and Globalization in Latin America*. Baltimore: Johns Hopkins University Press and Woodrow Wilson Center.

Hanushek, E.A. and M.E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24(2): 297-327.

Hanushek, E. A. and Luque. 2003. "Efficiency and Equity in Schools around the World." *Economics of Education Review* 20(5): 481-502.

Hoxby, C.M. 2002. "The Cost of Accountability," in W.M. Evers and H.J. Walberg, eds, *School Accountability*. Stanford: Hoover Institution Press.

Hoxby, C.M. 1996. "How Teachers' Unions Affect Education Production." *Quarterly Journal of Economics* 111(3): 671-718.

Jacob, B.A. 2005. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89(5-6):761-796.

Kasten, R. and G. Fossedal. n.d. "Teacher union "concentration" in 21 countries." Alexis de Tocqueville Institution (http://www.adti.net/gw-education.html).

Murillo, M.V. 1999. "Recovering Political Dynamics: Teachers' Unions and the Decentralization of Education in Argentina and Mexico." *Journal of Interamerican Studies and World Affairs* 41(1): 31-57.

Murillo, M.V., M. Tommasi, L. Ronconi and J. Sanguinetti. 2002. "The Economic Effects of Unions in Latin America: Teachers' Unions and Education in Argentina." Inter-American Development Bank, Latin American Research Network Working Paper R-463.

OECD. 2005. *Economic Survey of Mexico 2005*. Paris: OECD.

Ornelas, C. 2004. "The politics of privatisation, decentralisation and education reform in Mexico." *International Review of Education* 50(3-4): 397 – 418.

Ornelas, C. 1988. "The decentralization of education in Mexico." *Prospects* 18(1): 105-112.

Ornelas, C. n.d. "The Politics of the Educational Decentralization in Mexico." Autonomous Metropolitan University of Mexico (processed)..

Skoufias, E. and J. Shapiro. 2006. "Evaluating the impact of Mexico's quality schools program: the pitfalls of using nonexperimental data." World Bank Policy Research Working Paper Series No. 4036.

Woessmann, L. 2003. "Schooling Resources, Educational Institutions, and Student Performance: The International Evidence." *Oxford Bulletin of Economics and Statistics* 65(2): 117-170.

Wolter, S.C. and M.C. Vellacott. 2002. "Sibling Rivalry: A Look at Switzerland with PISA Data." IZA Discussion Papers No. 594, Bonn.

World Bank. 2005. *Mexico: Determinants of Learning Policy Note* (Report No. 31842-MX) Latin America and the Caribbean, Human Development.

World Bank. 2004. *World Development Report: Making Services for Poor People*. Washington DC: The World Bank.

Annex Table 1: Institutional Variables by State

	Government-Union conflict	Evaluation	Union influence on teacher positions	Within-state decentralization	
				Administrative	Pedagogical
Aguascalientes	3	4	1	no	yes
Baja California	3	4	1	yes	yes
Campeche	2	3	2	yes	no
Chiapas	3	3	2	no	no
Chihuahua	3	2	1	no	no
Coahuila	2	2	1	yes	no
Colima	3	5	1	yes	yes
Distrito Federal	2	5	2	yes	yes
Durango	3	2	2	yes	no
Guanajuato	3	4	1	yes	yes
Guerrero	1	1	3	yes	no
Hidalgo	2	4	2	yes	no
Jalisco	3	1	1	yes	no
México	2	1	1	no	no
Morelos	3	4	2	yes	no
Nayarit	2	2	2	no	no
Nuevo León	3	4	1	yes	yes
Oaxaca	1	2	3	yes	no
Puebla	3	1	2	yes	no
Querétaro	3	3	1	yes	yes
Quintana Roo	3	3	1	yes	no
San Luis Potosí	3	2	2	no	no
Sinaloa	3	2	2	no	no
Sonora	3	4	2	yes	no
Tabasco	2	1	1	no	no
Tamaulipas	3	2	2	no	no
Tlaxcala	1	1	3	no	no
Veracruz	2	2	2	no	no
Yucatán	3	2	1	no	no
Zacatecas	2	1	3	no	no

Note: Baja California Sur and Michoacán were not included in the analysis because of the lack of data

Annex Table 2: Institutional Effects as Determinants of Student's Math Achievement (PISA 2003)

	1		2		3		4		5	
Institutional factors	Coef.	S.E	Coef.	S.E	Coef.	S.E	Coef.	S.E	Coef.	S.E
Administrative decentralization (within states)	2.4	(1.99)***								
Pedagogical decentralization (within states)	3.8	(2.78)*								
Evaluation (second stage)			5.6	(1.50)*						
Evaluation (third stage)			-2.2	(2.02)						
Evaluation (fourth stage)			3.5	(1.60)*						
Evaluation (complete process and design strategies)			15.3	(2.33)*						
Parents' participation (medium)					2.2	(1.31)***				
Parents' participation (high)					0.8	(1.85)				
Medium power decision on teachers positions by union							-4.8	(2.17)*		
High power decision on teachers positions by union							-0.36	(2.16)		
Medium conflict between state ministry of education and union									1.2	(2.2)
Low conflict between state ministry of education and union									5.4	(2.1)*
Student characteristics										
Small City	18.1	(1.31)*	17.7	(1.30)*	18.2	(1.30)*	18.8	(1.44)*	17.7	(1.30)*
City	22.1	(1.63)*	21.8	(1.63)*	23.9	(1.57)*	25.8	(1.76)*	23.4	(1.56)*
Age	-1.2	(0.17)*	-1.2	(0.18)*	-1.1	(0.18)*	-1.1	(0.20)*	-1.1	(0.18)*
Female	-23.5	(1.13)*	-23.5	(1.13)*	-23.5	(1.13)*	-23.5	(1.25)*	-23.6	(1.12)*
Attitude	6.8	(0.56)*	6.72	(0.56)*	6.8	(0.57)*	6.8	(0.57)*	6.8	(0.57)*
Family Background										
Mother Education	10.5	(1.29)*	10.1	(1.29)*	10.4	(1.29)*	10.2	(1.42)*	10.3	(1.27)*
Mother working	2.1	(1.18)***	1.8	(1.18)	2	(1.18)**	1.8	(1.30)	1.8	(1.16)
Homework	2.5	(0.09)*	2.5	(0.09)*	2.5	(0.09)*	2.4	(0.10)*	2.5	(0.09)*
Home educational resources	8.7	(0.49)*	8.7	(0.49)*	8.7	(0.49)*	8.7	(0 .54)*	8.7	(0.49)*
Internet	-5.1	(0.36)*	-5.1	(0.36)*	-5.2	(0.36)*	-5.2	(0.36)*	-5.1	(0.36)*
Use of computer at home	2.4	(0.40)*	2.4	(0.40)*	2.5	(0.40)*	2.3	(0.40)*	2.4	(0.40)*
Motivation in Math	0.4	(0.76)	0.3	(0.76)	0.3	(0.76)	0.4	(0.83)	0.4	(0.76)
Memorization	-0.7	(0.63)	-0.6	(0.63)	-0.6	(0.63)	-0.5	(0.69)	-0.6	(0.63)
Teacher Morale	3.5	(0.51)*	3.7	(0.51)*	3.6	(0.5)*	3.5	(0.5)*	3.5	(0.50)*
Sense of belonging to school	0.6	(0.58)	0.5	(0.59)	0.6	(0.58)	0.8	(0.64)	0.7	(0.58)
Private School	2.3	(1.13)***	2	(1.13)***	2.1	(1.13)***	1.1	(1.26)	1.5	(1.12)
% of girls in the school	42.7	(4.57)*	44.4	(4.56)*	43.2	(4.58)*	41.7	(5.05)*	42.9	(4.47)*
Log Likelihood	-67,874		-67,853		-67,881		-70,325		-70,326	
Observations	12,281		12,281		12,281		12,719		12,719	

Source: Estimation with GLS form PISA 2003 and Mexican Institutional Variables

Annex Table 3: Institutional Effects as Determinants of Student's Achievement (PISA 2003)

Institutional factors	Math		Reading		Science		All	
	Coef.	S.E	Coef.	S.E	Coef.	S.E	Coef.	S.E
Administrative decentralization (within states)	0.4	(1.6)	0.04	(1.6)	1.7	(1.6)	0.7	(1.4)
Pedagogical decentralization (within states)	3.1	(2.8)	0.6	(2.9)	4.7	(2.8)	2.8	(2.5)
Evaluation (second stage)	2.0	(1.7)	-0.3	(1.8)	0.8	(1.7)	0.8	(1.5)
Evaluation (third stage)	-4.4	(2.3)	-1.2	(2.4)	-7.2	(2.3)*	-4.3	(2.1)*
Evaluation (fourth stage)	-2.1	(2.4)	-3.9	(2.5)	-5.4	(2.4)*	-3.8	(2.2)
Evaluation (complete process and design strategies)	14.7	(3.9)*	12.4	(4.0)*	6.1	(3.9)	11.1	(3.5)*
Medium power decision on teachers positions by union	-5.5	(3.0)	-5.4	(3.0)	-4.1	(3.0)	-5.0	(2.6)
High power decision on teachers positions by union	1.9	(3.2)	2.0	(3.2)	1.7	(3.2)	1.8	(2.8)
Medium conflict between state ministry of education and union	4.3	(2.5)	3.4	(2.5)	6.8	(2.5)*	6.8	(2.2)*
Low conflict between state ministry of education and union	9.2	(2.4)*	9.0	(2.5)*	9.3	(2.4)*	9.3	(2.1)*
Student characteristics								
Small City	17.4	(1.3)*	20.3	(1.4)*	15.2	(1.3)*	17.7	(1.2)*
City	21.7	(1.7)*	23.9	(1.7)*	20.2	(1.7)*	21.9	(1.5)*
Age	-1.2	(0.2)*	-1.2	(0.2)*	-1.5	(0.2)*	-1.3	(0.2)*
Female	-22.9	(1.1)*	10.9	(1.1)*	-21.7	(1.1)*	-11.2	(1.0)*
Family Background								
Mother Education	10.1	(1.3)*	13.42	(1.3)*	13.35	(1.3)*	12.3	(1.1)*
Mother working	1.8	(1.2)	0.6	(1.2)	2.3	(1.2)	1.54	(1.0)
Home incentives and inputs								
Homework	2.6	(0.1)*	2.5	(0.1)*	2.8	(0.1)*	2.7	(0.1)*
Home educational resources	9.1	(0.5)*	9.84	(0.5)*	8.79	(0.5)*	9.2	(0.5)*
Internet	-5.1	(0.4)*	-4.9	(0.4)*	-5.5	(0.4)*	-5.1	(0.3)*
Use of computer at home	2.5	(0.4)*	2.3	(0.4)*	3.0	(0.4)*	2.60	(0.4)*
Schools resources								
Memorization	0.1	(0.6)	-3.0	(0.6)*	0.8	(0.6)*	-0.7	(0.5)*
Teacher Morale	3.8	(0.5)*	4.7	(0.5)*	3.5	(0.5)*	4.0	(0.5)*
Sense of belonging to school	2.4	(0.6)*	4.3	(0.6)*	2.32	(0.6)*	3.0	(0.5)*
Private School	2.0	(1.1)	2.6	(1.2)*	2.3	(1.1)	2.3	(1.0)
% of girls in the school	45.2	(4.6)*	42.7	(4.6)*	34.8	(4.6)*	40.9	(4.1)*
Log Likelihood	-68,188		-68,349		-68,152		-66,727	
Observations	12,332		12,332		12,332		12,332	

Source: Estimation with GLS form PISA 2003 and Mexican Institutional Variables
* Denotes significance at the 95% level

Annex 4: Institutional Effects as Determinants of Student Achievement using Propensity Matching Score: Simulated Scores

Institutional factors	**Math** ATT -Diff.	**Read** ATT -Diff.	**Science** ATT -Diff.	**All** ATT -Diff.
Evaluation (1st stage vs full accountability)	28.2 (3.85)	28.3 (3.9)	21.8 (3.76)	26.1 (3.41)
Evaluation (3rd stage vs full accountability)	18.1 (3.45)	16.8 (3.43)	11.6 (3.31)	15.5 (3.41)

Notes: ATT is the average treatment on the treated; Diff. refers to the difference between treated and control scores. Full results available upon request. Standard Errors in parenthesis

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS4263	HIV/AIDS and Social Capital in a Cross-Section of Countries	Antonio C. David	June 2007	A. David 82842
WPS4264	Financing of the Private Sector in Mexico, 2000–05: Evolution, Composition, and Determinants	Constantinos Stephanou Emanuel Salinas Muñoz	June 2007	S. Coca 37474
WPS4265	The Structure of Import Tariffs in the Russian Federation: 2001–05	Oleksandr Shepotylo	June 2007	P. Flewitt 32724
WPS4266	The Economic Community of West African States: Fiscal Revenue Implications of the Prospective Economic Partnership Agreement with the European Union	Simplice G. Zouhon-Bi Lynge Nielsen	June 2007	S. Zouhon-Bi 82929
WPS4267	Financial Intermediation in the Pre-Consolicated Banking Sector in Nigeria	Heiko Hesse	June 2007	G. Johnson 34436
WPS4268	Power to the People: Evidence from a Randomized Field Experiment of a Community-Based Monitoring Project in Uganda	Martina Björkman Jakob Svensson	June 2007	I. Hafiz 37851
WPS4269	Shadow Sovereign Ratings for Unrated Developing Countries	Dilip Ratha Prabal De Sanket Mohapatra	June 2007	N. Aliyeva 80524
WPS4270	Jump-Starting Self-Employment? Evidence among Welfare Participants in Argentina	Rita Almeida Emanuela Galasso	June 2007	A. Bonfield 31248
WPS4271	Construction, Corruption, and Developing Countries	Charles Kenny	June 2007	C. Kenny 33540
WPS4272	Migration, Remittances, Poverty, and Human Capital: Conceptual and Empirical Challenges	David McKenzie Marcin J. Sasin	July 2007	M. Sasin 36877
WPS4273	Rules of Origin and the Web of East Asian Free Trade Agreements	Miriam Manchin Annette O. Pelkmans-Balaoing	July 2007	L. Yeargin 81553
WPS4274	Are Labor Regulations Driving Computer Usage in India's Retail Stores?	Mohammad Amin	July 2007	S. Narsiah 88768
WPS4275	Can Foreign Lobbying Enhance Development? The Case of Tourism in the Caribbean	Kishore Gawande William Maloney Gabriel V. Montes Rojas	July 2007	V. Cornago 84039
WPS4276	Human Capital, Trade Liberalization, and Income Risk	Tom Krebs Pravin Krishna William Maloney	July 2007	V. Cornago 84039
WPS4277	Climate Change Adaptation in Africa: A Microeconomic Analysis of Livestock Choice	Sungno Niggol Seo Robert Mendelsohn	July 2007	P. Kokila 33716
WPS4278	Endogenous Irrigation: The Impact of Climate Change on Farmers in Africa	Pradeep Kurukulasuriya Robert Mendelsohn	July 2007	P. Kokila 33716
WPS4279	The Impact of Climate Change on Livestock Management in Africa: A Structural Ricardian Analysis	Sungno Niggol Seo Robert Mendelsohn	July 2007	P. Kokila 33716

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS4280	Governance Matters VI: Aggregate and Individual Governance: Indicators, 1996-2006	Daniel Kaufmann Aart Kraay Massimo Mastruzzi	July 2007	R. Bonfield 31248
WPS4281	Credit Growth In Emerging Europe: A Cause For Stability Concerns?	Sophie Sirtaine Ilias Skamnelos	July 2007	S. Sirtaine 87006
WPS4282	Are Cash Transfers Made to Women Spent Like Other Sources of Income	Norbert Schady José Rosero	July 2007	I. Hafiz 37851
WPS4283	Innovation Shortfalls	William Maloney Andrés Rodríguez-Clare	July 2007	V. Cornago 84039
WPS4284	Customer Market Power and the Provision of Trade Credit: Evidence from Eastern Europe and Central Asia	Neeltje Van Horen	July 2007	M. Gamboa 34847
WPS4285	Poverty Analysis Using An International Cross-Country Demand System	J. A. L. Cranfield Paul V. Preckel Thomas W. Hertel	July 2007	P. Flewitt 32724